

mmIR: Frequency-Space Inverse Rendering for 3D Millimeter-Wave Radar ADC Synthesis

Adnan Armouti¹, Yixuan Gao¹, and Rajalakshmi Nandakumar¹

Cornell Tech, New York, NY, USA
{aa2546,yg478,rn283}@cornell.edu

Abstract. High-resolution 3D radar data is scarce. Commodity mmWave sensors use small antenna arrays that limit angular resolution to several degrees, and existing datasets provide only 2D range–azimuth maps or sparse point clouds rather than raw analog-to-digital converter (ADC) signals. Hardware scaling is expensive, synthetic-aperture scanning is impractical at fleet scale, and learned synthesis methods are bottlenecked by the very data shortage they aim to address. We present mmIR, an open-source differentiable frequency-modulated continuous-wave (FMCW) radar inverse renderer that fits a physics-based forward model to real captures and re-renders from dense virtual apertures to synthesize high-resolution 3D radar data. Because radar resolution is too coarse to recover geometry directly, mmIR performs LiDAR-assisted inverse rendering: using LiDAR-derived meshes as a geometric scaffold, mmIR optimizes per-vertex International Telecommunication Union (ITU) physics materials, vertex normals, and antenna beam patterns through end-to-end automatic differentiation of a phase-coherent multiple-input multiple-output (MIMO) forward model with multi-bounce propagation, polarization, and free-space diffraction. On seven outdoor and six indoor ColoRadar scenes, mmIR achieves 0.914 mean Pearson correlation on range–azimuth maps versus 0.307 for Sionna-RT. Scenes trained on a cascaded imaging radar transfer to a co-located single-chip radar without re-training (0.554 correlation), and dense virtual arrays (100×100 elements) produce single-frame 3D occupancy validated against LiDAR. Project page: <https://mmwave-inverse-rendering.github.io/>.

Keywords: Differentiable rendering · FMCW radar · Inverse rendering · mmWave · 3D reconstruction

1 Introduction

Millimeter-wave (mmWave) radar penetrates fog, rain, and airborne dust [16], measures Doppler velocity, and returns strong echoes from metals, yet it remains underutilized for 3D scene understanding. Commercial single-chip radars have only tens of antennas, limiting angular resolution to several degrees [14], and their axis-biased layouts yield 2D range–azimuth (RA) slices rather than full 3D range–azimuth–elevation (RAE) volumes. Public datasets further limit access by providing only post-processed outputs (RA maps or sparse point clouds)

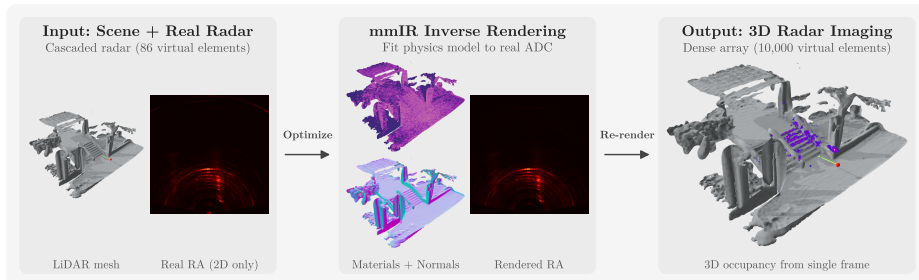


Fig. 1: Given a LiDAR mesh and GT real cascaded-radar RA map (left), mmIR fits per-vertex materials, normals, and beam patterns (center), then re-renders a dense 100×100 virtual array for single-frame 3D occupancy (right).

rather than raw ADC [4, 9]. The result is a shortage of high-resolution, 3D, raw FMCW ADC data, the fundamental bottleneck for learning-based radar perception. Detection, segmentation, occupancy prediction, and super-resolution all require training signal that current sensors and datasets cannot supply, and addressing this gap would enable these applications.

Existing approaches each fall short. Hardware scaling [42] increases cost without addressing data diversity. SAR [31, 50, 56] requires mechanical scanning impractical at fleet scale. Neural and Gaussian synthesis methods [8, 19, 27, 36, 43] are bottlenecked by the very data scarcity they aim to address. Models trained on low-resolution 2D RA maps cannot recover 3D structure that was never observed.

Physics-based ray tracing (SBR) for mmWave is well established [13, 18], but existing pipelines are forward-only, requiring geometry and materials to be specified *a priori*, with no mechanism to calibrate against real measurements. Differentiable rendering has closed this gap for RGB [29, 55], LiDAR [23], and acoustics [22], and emerging RF methods take initial steps. Sionna-RT [18] exposes gradients for antenna patterns and scattering coefficients but not through ray tracing itself (no geometry or normal gradients) and does not model wideband FMCW signals. Hofmann et al. [17] estimate SFCW materials but assume known geometry with single-bounce propagation, and InverTwin [10] addresses multipath gradient pathologies but validates on synthetic data only. What is missing is an end-to-end differentiable FMCW renderer that can learn scene parameters from real radar captures without requiring ground-truth materials.

We present mmIR, an open-source differentiable FMCW radar inverse renderer that fits a physics-based forward model to real captures and re-renders from dense virtual apertures to synthesize high-resolution 3D radar data. Radar resolution is too coarse to reconstruct geometry directly, so mmIR adopts a LiDAR-assisted formulation, using LiDAR-derived meshes [24] as a geometric scaffold, a practical choice given the far greater availability and resolution of LiDAR data. Inverse rendering then recovers the electromagnetic properties. mmIR optimizes per-vertex ITU physics materials (permittivity, roughness, correlation length, blend factor, slab thickness), vertex normals, and antenna beam patterns through end-

to-end automatic differentiation of a phase-coherent MIMO forward model with multi-bounce propagation, polarization-aware bidirectional scattering distribution functions (BSDFs), specular manifold sampling [53], and free-space diffraction [41]. The pipeline also supports differentiable geometry and 6-DoF pose gradients with projective boundary sampling [57], though we keep these fixed in our experiments because LiDAR-radar alignment preprocessing (Sec. B.2 in the supplement) already provides accurate geometry and pose. After fitting, the optimized scene is frozen and re-rendered from arbitrarily large virtual apertures to synthesize dense, physics-consistent raw ADC (Fig. 1), enabling downstream applications such as 3D occupancy detection, radar super-resolution, and synthetic training data generation that would otherwise require prohibitively large physical arrays.

On seven outdoor and six indoor ColoRadar [26] scenes, mmIR achieves 0.914 mean Pearson RA correlation versus 0.307 for Sionna-RT. We demonstrate cross-sensor transfer. Scenes trained on a cascaded imaging radar (12 transmit (Tx) $\times$ 16 receive (Rx) antennas spanning 86 uniformly sampled virtual elements along azimuth) produce valid ADC for a co-located single-chip radar (3Tx $\times$ 4Rx, 12 virtual elements total spanning 8 uniformly sampled virtual elements along azimuth) without re-training (0.554 correlation), providing evidence that the optimized scenes capture transferable physics rather than sensor-specific patterns. We further show single-frame 3D occupancy detection from dense virtual arrays (100Tx $\times$ 100Rx) validated against LiDAR. mmIR’s scene fitting takes ~ 11 min (500 iterations, 1.32 s/iter) versus 3.28 s/iter for Sionna-RT (Sec. C.5), and a dense-array frame renders in 6.4 s on a single NVIDIA RTX 4090. Our contributions are:

1. To the best of our knowledge, the first open-source, end-to-end differentiable inverse renderer for wideband FMCW radar, enabling joint optimization of per-vertex materials, normals, and antenna patterns directly from real radar captures, with support for multi-bounce propagation, specular manifold sampling, polarization-aware BSDFs, free-space diffraction, and differentiable geometry and pose gradients with projective boundary sampling.
2. A virtual-aperture re-rendering framework that synthesizes high-resolution 3D raw ADC from fitted scenes via arbitrarily large synthetic arrays, enabling radar resolution lifting from 2D RA to 3D RAE on the provided LiDAR scaffold, and 3D occupancy detection.
3. Cross-sensor transfer validation on real COTS radar, demonstrating that scenes optimized on one sensor generalize to a different configuration without re-training.

1.1 Scope

mmIR targets high-quality, physics-consistent raw-ADC synthesis by fitting a differentiable forward model on LiDAR-derived geometry to real commercial off-the-shelf (COTS) radar measurements. We envision practitioners using optimized scenes to build large-scale datasets for training downstream perception models such as detection, segmentation, and occupancy prediction, without additional

Table 1: Feature comparison of physics-based differentiable radar renderers. Signal: ADC = raw FMCW, CIR = channel impulse response, IF = intermediate-frequency phasors. Multi-b. = multi-bounce; Diff. = edge diffraction; Polar. = polarization; Geom. = differentiable geometry; Norm. = differentiable normals; Mat. = physics material model; Per-vtx = per-vertex materials; Bound. ∇ = boundary gradients; Real = validated on real data. Checkmarks denote supported pipeline capabilities; see Sec. B.4 in the supplement for which parameters are active in our experiments.

Method	Open	Repr.	Signal	Multi-b.	Diff.	Polar.	Geom.	Norm.	Mat.	Pose	Beam	Per-vtx	Bound. ∇	Real
Sionna-RT [18]	✓	Mesh	CIR	✓	✓	✓	–	–	No BRDF	✓	–	–	–	–
SFCW Inv. [17]	✓	Mesh	IF	–	–	✓	–	Shading	✓	Pos.	–	✓	–	✓
InverTwin [10]	–	Mesh	CIR	✓	✓	–	✓	–	No BRDF	✓	–	–	✓	✓
mmIR (Ours)	✓	Mesh	ADC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

data collection. We show cross-sensor transfer and single-frame 3D occupancy from dense virtual arrays (Sec. 4.3), comparing only against open-source physics-based ray-tracing renderers (Table 1) for a reproducible, like-for-like evaluation. We also note that single-frame radar inverse rendering is severely ill-posed. Our experiments are run on a single view with a single frame, which suffers from sparse returns, elevation integrated out by 2D RA, and resolution far coarser than LiDAR. Therefore, we adopt a mesh-based geometric prior following Sionna-RT [18], but validate on meshes obtained from real LiDAR data rather than synthetic CAD, and fit to real radar measurements instead of synthetic channel responses. We do not claim shape-from-radar, an open problem under single-frame supervision that is out of scope. Finally, we assume static captures, leaving Doppler (ego-motion) and dynamic scenes to future work.

2 Related Works

Differentiable rendering beyond optics. Differentiable rendering couples explicit scene parameterizations with gradient-based optimization to recover geometry, materials, and sensor parameters [15, 39, 46, 47]. Core advances include analytic inverse rendering and Monte Carlo (MC) estimator reparameterizations [2, 29, 33], as well as practical path tracing systems and neural/hybrid representations [1, 3, 21, 35, 37, 55]. These methods operate almost exclusively in the optical regime. In RF, Sionna-RT [18] propagates gradients through field coefficients and CIR computation but not through path topology. Its path solver runs under gradient suspension and subsequently detaches all path geometry, so vertex positions and surface normals receive no gradients, restricting inverse rendering to quantities that enter only the CIR computation (antenna patterns, scattering coefficients, Tx/Rx positions). It also does not model wideband FMCW signals. Hofmann et al. [17] estimate SFCW materials but require known geometry, model only single-bounce propagation, and treat visibility as non-differentiable. InverTwin [10] addresses multipath gradient pathologies via path-space differentiation but substitutes a Gaussian kernel surrogate for radar range processing and validates on synthetic data only. mmIR bridges this gap with end-to-end AD through ray tracing itself, enabling gradients to flow through all bounces back to

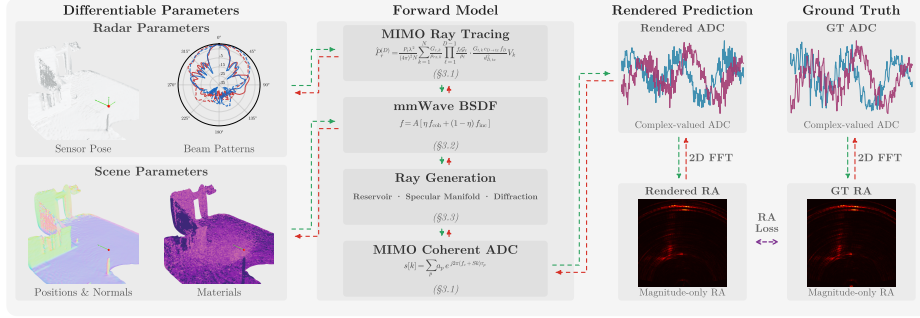


Fig. 2: mmIR pipeline. Given a LiDAR mesh and sensor parameters, the differentiable model traces phase-coherent MIMO paths (reflection + diffraction) into raw FMCW ADC. End-to-end AD (ray tracing, BSDF, phasor accumulation, ADC) optimizes per-vertex materials, normals, and antenna beam patterns from a range-azimuth loss.

geometry, materials, and sensor parameters directly from raw frequency-space FMCW ADC (Table 1).

Ray tracing for FMCW signal modeling. Radar simulation typically employs shooting-and-bouncing rays (SBR) to scale beyond full-wave solvers, producing path geometry and attenuation for large scenes and MIMO arrays [13]. Recent computational RF imaging further demonstrates multi-bounce physics for reconstruction on real mmWave hardware [12]. However, these pipelines are forward-only. Scene and sensor parameters are fixed, precluding calibration against real captures. mmIR makes SBR-style FMCW rendering fully differentiable, exposing gradients w.r.t. mesh geometry, per-vertex materials, radar pose, and antenna patterns to close the sim-to-real gap.

Neural and Gaussian methods for radar data synthesis. Neural fields have been adapted to radar across modalities, including SAR/ISAR shape reconstruction [11, 28], FMCW occupancy and Doppler recovery [19, 43], frequency-space response synthesis [8, 54], and multi-modal fusion with RGB/LiDAR [34, 38]. RadarSplat [27] provides a Gaussian splatting [25] alternative for range-azimuth rendering. These methods learn opaque, sensor-specific representations. DART [19] discards phase entirely (magnitude-only range-Doppler supervision), none expose physically interpretable material parameters, and none model multi-bounce propagation, polarization, or diffraction. Moreover, methods that render post-processed images directly must learn to reproduce FFT point-spread artifacts (sidelobes, windowing). By rendering raw ADC, mmIR applies identical FFT processing to both rendered and measured data, cleanly decoupling scene physics from signal processing. mmIR maintains an explicit mesh with per-vertex ITU materials and polarization-aware BSDFs, enabling joint scene-sensor calibration and re-rendering from arbitrary virtual apertures for 3D imaging.

3 Methods

We model radar sensing as a forward rendering problem (Fig. 2), tracing multi-bounce rays through a scene, evaluating surface interactions via a physically-based mmWave BSDF, and accumulating phase-coherent paths into raw FMCW ADC samples that can be directly compared against real captures. Sec. 3.1 derives the forward model from the bistatic radar equation through multi-bounce Monte Carlo estimation to ADC synthesis; Sec. 3.2 defines the mmWave BSDF governing surface interactions; Sec. 3.3 describes ray generation strategies; and Sec. 3.4 formulates the differentiable inverse rendering pipeline.

Our inputs are synchronized radar–LiDAR data from ColoRadar [26]. LiDAR frames are aggregated into a triangle mesh via screened Poisson reconstruction [24], and a GPU-accelerated coarse-to-fine alignment corrects residual sensor misregistration (Sec. B.1–B.2 in the supplement). FMCW signal and TDM-MIMO imaging derivations are in the supplement (Sec. A–A.3).

3.1 Forward Model

Our forward model is grounded in the bistatic radar equation. For a single Tx–Rx pair observing a point target at surface point $\mathbf{x}$ with normal $\mathbf{n}$, at distances d_t , d_r from transmitter and receiver, the received power P_r is:

$$P_r = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 d_t^2 d_r^2}, \quad (1)$$

where P_t is transmit power, G_t , G_r are directional antenna gains, λ is wavelength, and σ is the bistatic radar cross-section. Replacing the point-target RCS with a spatially varying BSDF $f(\boldsymbol{\omega}_t, \mathbf{x}, \boldsymbol{\omega}_r)$, where $\boldsymbol{\omega}_t$ and $\boldsymbol{\omega}_r$ are unit directions toward the transmitter and receiver, and converting from surface area element dA_x to receiver solid angle via $d\omega_r = (c_r/d_r^2) dA_x$ with $c_r = \max(0, \mathbf{n} \cdot \boldsymbol{\omega}_r)$ (derivation in the supplement, Sec. A.4), the total received power integrated over the receiver hemisphere Ω_{rx} becomes:

$$P_r = \frac{P_t \lambda^2}{(4\pi)^2} \int_{\Omega_{rx}} \frac{G_t(\boldsymbol{\omega}_t) G_r(\boldsymbol{\omega}_r) f(\boldsymbol{\omega}_t, \mathbf{x}, \boldsymbol{\omega}_r) c_t V}{d_t^2} d\omega_r, \quad (2)$$

where $c_t = \max(0, \mathbf{n} \cdot \boldsymbol{\omega}_t)$ is the cosine factor and $V \in \{0, 1\}$ is visibility. This solid-angle formulation directly motivates Rx-centric sampling: trace rays from the receiver, intersect the scene, and evaluate the integrand with a deterministic connection to the transmitter.

For multi-bounce propagation with path depth D and interaction points $\mathbf{x}_1, \dots, \mathbf{x}_D$, each stochastically sampled segment contributes a BSDF-over-PDF ratio and geometric coupling. Given N Monte Carlo samples, the unbiased estimator is (derivation in the supplement, Sec. A.4):

$$\hat{P}_r^{(D)} = \frac{P_t \lambda^2}{(4\pi)^2} \frac{1}{N} \sum_{k=1}^N \frac{G_{r,k}}{\rho_{rx,k}} \prod_{\ell=1}^{D-1} \frac{f_\ell \mathcal{G}_\ell}{\rho_\ell} \cdot \frac{G_{t,k} c_{D \rightarrow tx} f_D}{d_{D,tx}^2} V_k, \quad (3)$$

where $\mathcal{G}_\ell = c_{\ell \rightarrow \ell+1} c_{\ell+1 \rightarrow \ell} / d_{\ell, \ell+1}^2$ is the geometric coupling between consecutive bounces, with $c_{\ell \rightarrow \ell+1} = \max(0, \mathbf{n}_\ell \cdot \boldsymbol{\omega}_{\ell \rightarrow \ell+1})$ the cosine factor at bounce ℓ toward bounce $\ell+1$ and $d_{\ell, \ell+1}$ the distance between them. The BSDF at each bounce is $f_\ell = f(\boldsymbol{\omega}_{\ell-1}, \mathbf{x}_\ell, \boldsymbol{\omega}_{\ell+1})$ (Sec. 3.2), which decomposes each surface interaction into slab transmission, coherent specular reflection, incoherent diffuse scattering, and edge diffraction. The final term connects deterministically to the transmitter: $c_{D \rightarrow tx}$ is the cosine factor at $\mathbf{x}_D$ toward the Tx and $d_{D, tx}$ is their distance. The sampling PDFs ρ_ℓ are determined by the ray generation strategies in Sec. 3.3. The single-bounce case ($D=1$) reduces to $G_t G_r c_t f V / (d_t^2 \rho_{rx})$, and we use Russian roulette [46] for unbiased path termination.

The MC estimator provides a complex amplitude for each traced path. To synthesize raw FMCW ADC, each path p for Tx i and Rx j contributes a phase-delayed chirp sample using the complex baseband form (Sec. A.1–A.2 in the supplement):

$$s_{ij}^{(p)}[k] = a_{ij}^{(p)} \exp(j 2\pi(f_c + S t_k) \tau_{ij}^{(p)}), \quad (4)$$

where k indexes fast-time ADC samples at times t_k , j is the imaginary unit, $\tau_{ij}^{(p)} = R_{ij}^{(p)} / c$ is the round-trip delay for total path length $R_{ij}^{(p)}$ at speed of light c , f_c is the carrier frequency, S is the chirp slope, and $a_{ij}^{(p)} \in \mathbb{C}$ aggregates the MC throughput from Eq. (3), namely antenna gains, BSDF evaluations, geometric coupling, and spreading losses. The rendered ADC sums coherently over all paths: $s_{ij}[k] = \sum_p s_{ij}^{(p)}[k]$.

Equations (1)–(4) describe a single Tx–Rx pair. In a MIMO array, each pair is evaluated independently through the same forward model. However, coherent MIMO imaging via FFT [30] requires exact relative phase across all pairs, so independent MC samples per receiver would corrupt this. Following the MIMO ray reuse strategy of Schüßler et al. [40], we share path topology across all Tx–Rx pairs. A single set of interaction points is found from shared seeds, then evaluated for all pairs by vectorizing only the Tx $\rightarrow \mathbf{x}_1$ and $\mathbf{x}_D \rightarrow$ Rx legs. Each pair produces unique amplitudes and phases but shares bounce topology, preserving MIMO coherence while reducing cost from $\mathcal{O}(N_{\text{Tx}} \cdot N)$ to $\mathcal{O}(N)$.

3.2 Physically-Based mmWave BSDF

The BSDF f appearing in the MC estimator (Eq. (3)) governs how electromagnetic energy is redistributed at each surface interaction. We model it using ITU-R P.2040 recommendations for mmWave propagation [20], inspired by the coherent spatially varying BSDF (CSVBSDF) formulation of Wei et al. [51]. Each mesh vertex stores six physically interpretable material parameters, namely real and imaginary relative permittivity $(\epsilon'_r, \epsilon''_r)$, RMS surface roughness σ_h , correlation length l_c , a Kirchhoff-approximation / small-perturbation-method (KA/SPM) blend factor τ , and slab thickness d . At each ray–triangle intersection, parameters are barycentric-interpolated from the three triangle vertices $(\mathbf{m}_i \in \mathbb{R}^6)$, ensuring C^0 -continuous material fields with gradients flowing to all three vertices.

The BSDF decomposes into coherent and incoherent components via the Rayleigh roughness factor [5] $\eta = \exp(-(2k\sigma_h \cos \theta_t)^2)$, $k = 2\pi/\lambda$, where θ_t is the incidence angle between ω_t and the surface normal:

$$f = A(\omega_t) [\eta f_{\text{coh}} + (1 - \eta) f_{\text{inc}}], \quad (5)$$

where $A(\omega_t) \in [0, 1]$ is the slab Fresnel energy gate. The lobe decompositions are:

$$f_{\text{coh}} = \tau f_{\text{KA}} + (1 - \tau) f_{\text{SPM}}, \quad (6)$$

$$f_{\text{inc}} = \gamma f_{\text{dir}} + (1 - \gamma) f_{\text{broad}}, \quad (7)$$

where τ is a learnable KA/SPM blend and γ is a roughness-dependent sigmoid. The coherent lobes are f_{KA} and f_{SPM} . f_{KA} is a GGX microfacet specular lobe [49] with roughness $\alpha = \sqrt{4\pi\sigma_h/\lambda}$, and f_{SPM} is a von Mises–Fisher (vMF) distribution centered on the specular direction with concentration $\kappa = \sqrt{l_c/\lambda} / (1 + 3\sigma_h/l_c)$ [45]. The incoherent lobes are f_{dir} , a broader vMF lobe, and f_{broad} , Lambertian. The energy gate $A(\omega_t)$ uses an ITU-R P.2040 slab model with thickness d and complex permittivity $\varepsilon_r = \varepsilon'_r - j\varepsilon''_r$, capturing multiple internal reflections rather than single-interface Fresnel.

The Fresnel coefficients in the energy gate also determine polarization state, which we maintain throughout path tracing using a local s/p basis. At each interaction, the complex Fresnel coefficients derived from learnable permittivity update both amplitude and phase of the s and p field components [7]. Between bounces, the polarization basis is transported deterministically to preserve phase consistency. The polarization-dependent energy gate is $A = f_s |r_s|^2 + f_p |r_p|^2$, where r_s, r_p are the complex Fresnel reflection coefficients for the s and p polarizations and f_s, f_p are the antenna-dependent power fractions in each polarization.

3.3 Ray Generation

The MC estimator (Eq. (3)) requires sampling ray directions at each interaction. Because the mmWave BSDF (Sec. 3.2) combines smooth diffuse, near-delta specular, and edge-diffraction components that no single sampling PDF can efficiently cover, we provide the PDFs ρ_ℓ via three complementary strategies, multiple importance sampling (MIS) [47] and balance heuristic weights.

Rx-centric reservoir sampling. Because mmWave antennas are highly directional, we employ Rx-centric reservoir sampling inspired by Resampled Importance Sampling (RIS) [44] and spatiotemporal reservoir resampling (ReSTIR) [6, 32]. For each Rx element, we sample M candidate directions from a cosine-weighted hemisphere proposal centered on the Rx boresight, $\rho_{rx}(\omega) = \cos \theta/\pi$, which naturally concentrates samples within the antenna main lobe, then retain K directions that successfully intersect scene geometry.

Specular manifold sampling. LiDAR-derived meshes contain triangles much smaller than the mmWave wavelength ($\lambda \approx 3.9$ mm). The image method (mirroring Tx across the hit triangle’s plane) frequently fails because the exact specular point falls outside the original triangle. We instead use Specular Manifold Sampling (SMS) [53], which solves for exact reflection points via Newton iteration on the half-vector constraint $\mathbf{C}(\mathbf{x}) = [\mathbf{s} \cdot \mathbf{h}, \mathbf{t} \cdot \mathbf{h}]^\top = \mathbf{0}$, where $\mathbf{s}, \mathbf{t}$ are surface tangents and $\mathbf{h}$ is the half-vector. After each Newton step, re-projection onto the mesh via ray tracing allows the solver to cross triangle boundaries, producing exact specular paths with deterministic Fresnel weights.

We extend SMS for radar inverse rendering with GPU-vectorized Newton iteration via DrJit [21], Implicit Function Theorem gradient support that stores the Jacobian inverse at convergence for backpropagation through vertex positions, and multi-bounce specular chain solving via a block-tridiagonal algorithm.

Free-space diffraction BSDF. Unlike the Uniform Theory of Diffraction (UTD), which requires explicit edge detection and separate ray classes, the free-space diffraction (FSD) BSDF [41] formulates edge diffraction as a standard scattering distribution compatible with path tracing. At each intersection, the FSD BSDF projects nearby triangles onto a virtual screen, identifies boundary edges, and evaluates closed-form Fraunhofer diffraction integrals. A fraction β of reflected energy is redirected into importance-sampleable diffracted lobes.

We extend FSD for mmWave radar with: (i) material-dependent edge opacity via Fresnel reflectance from learnable permittivity; (ii) Jones polarization tracking through diffracted paths for MIMO phase coherence; and (iii) edge suppression filters that reject artifacts from noisy LiDAR tessellation.

3.4 Differentiable Optimization

We optimize per-vertex ITU physics materials ($6 \times N_v$ parameters with sigmoid/exponential reparameterizations for physical bounds), per-vertex normals $\mathbf{N} \in \mathbb{R}^{N_v \times 3}$, and antenna beam patterns. The pipeline additionally supports differentiable vertex positions and 6-DoF radar pose, but we keep these fixed in our experiments, since our coarse-to-fine LiDAR-radar alignment (Sec. B.2 in the supplement) already provides sub-centimeter translational and sub-degree rotational accuracy, so the remaining sim-to-real gap is dominated by material and scattering properties rather than geometric misregistration. Vertex-position and pose optimization remain available as a fallback when the input mesh is degraded or alignment is poor (Sec. B.4, and limitations in Sec. D). Initialization uses ITU-R P.2040 reference values for common materials (concrete, glass, metal). Details on bounds, learning rates, and schedules are in the supplement (Sec. B.4).

The entire forward model (Secs. 3.1–3.3) executes inside a single DrJit [21] computation graph with automatic differentiation. Gradients flow from $\mathcal{L}$ through the ADC, per-path phasors, BSDF evaluations, and ray-triangle intersections back to all learnable parameters in a single backward pass. At silhouette edges where visibility changes discontinuously, projective boundary gradient sampling [57] corrects gradient bias for vertex positions and pose when these are enabled.

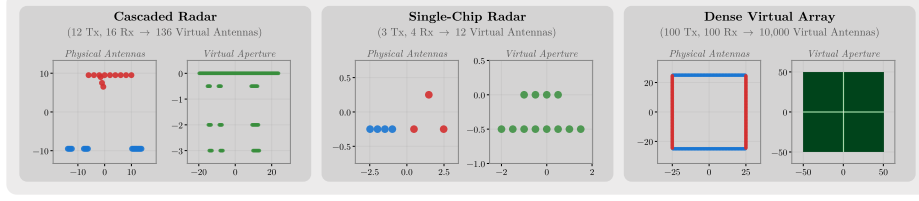


Fig. 3: Antenna configurations: cascaded MMWCAS (12Tx×16Rx with 136 unique virtual elements) used for training, single-chip AWR1843 (3Tx×4Rx with 12 unique virtual elements) used for transfer evaluation, dense array (100Tx×100Rx, 10,000 virtual elements) used for 3D inference. Top: physical Tx/Rx; bottom: virtual aperture. The cascade’s near-1D virtual aperture yields 256×127×1 RAE resolution (Sec. A.1, A.3); only the dense array resolves true elevation at 256×127×127 RAE resolution.

Loss function. We minimize a range–azimuth reconstruction loss on min–max normalized magnitudes:

$$\mathcal{L} = \frac{1}{|\Gamma|} \sum_{(r,a) \in \Gamma} \|\widetilde{\text{RA}}_{\text{render}}(r,a) - \widetilde{\text{RA}}_{\text{gt}}(r,a)\|_2^2, \quad (8)$$

where Γ is the set of range–azimuth bins, $\widetilde{\text{RA}} = (\text{RA} - \min \text{RA}) / (\max \text{RA} - \min \text{RA})$ denotes min–max normalization, and RA maps are obtained via the TDM-MIMO FFT pipeline [30]. RA magnitude validates phase coherence because the azimuth FFT encodes both relative phase across virtual elements and ADC magnitude. This is ideal, because the absolute phase depends on user-configurable initial phase, oscillator drift, and sub-millimeter path-length variation at 77 GHz and carries no reliable supervisory signal (Sec. B.3). After convergence, we freeze optimized parameters and re-render from dense virtual apertures to synthesize 3D radar data (Sec. 4.3).

4 Experiments

We evaluate mmIR on seven diverse outdoor scenes from the ColoRadar dataset [26], which provides synchronized mmWave radar and LiDAR captures with raw ADC access. We design three experiments that progressively test our forward model under increasingly challenging conditions: (1) training evaluation on the cascaded imaging radar used for optimization (12Tx×16Rx with 136 unique virtual elements, spanning 86 uniformly sampled virtual elements along the azimuth); (2) cross-sensor transfer to a co-located single-chip radar (3Tx×4Rx with 12 virtual elements, spanning 8 uniformly sampled virtual elements along the azimuth) without re-training; and (3) 3D occupancy inference via a dense virtual array (100Tx×100Rx with 10,000 unique virtual elements, spanning 100 uniformly sampled virtual elements along the azimuth and the elevation each). Figure 3 illustrates all three radar configurations and antenna layouts. Dataset details, scene geometry reconstruction, and implementation hyperparameters are provided in the supplement (Sec. B).

Table 2: Range–azimuth (RA) and ADC training evaluation on seven outdoor and six indoor scenes. mmIR is compared against Sionna-RT [18]. Metrics: Pearson correlation (Corr), PSNR, SSIM, RMSE, and ADC-level magnitude error. Best in **bold**. Results use the min–max-normalized RA loss (Eq. 8).

Scene	RA Correlation $\uparrow$		RA PSNR $\uparrow$		RA SSIM $\uparrow$		RA RMSE $\downarrow$		ADC Log-Mag MSE $\downarrow$		ADC MM-Mag MSE $\downarrow$	
	Ours	Sionna-RT	Ours	Sionna-RT	Ours	Sionna-RT	Ours	Sionna-RT	Ours	Sionna-RT	Ours	Sionna-RT
<i>Outdoor (7 scenes)</i>												
S0-F135	0.921	0.418	37.4	21.4	0.938	0.309	0.0135	0.0850	48.3271	244.8361	0.0421	0.0469
S0-F390	0.936	0.242	41.0	23.4	0.937	0.350	0.0089	0.0674	5.0250	294.3032	0.0289	0.0417
S1-F185	0.954	0.312	39.8	18.9	0.926	0.326	0.0102	0.1138	32.3385	282.6453	0.0398	0.0461
S1-F438	0.942	0.346	37.7	19.9	0.926	0.401	0.0130	0.1008	32.5274	288.9782	0.0369	0.0435
S2-F105	0.887	0.188	34.6	28.7	0.894	0.724	0.0185	0.0367	23.5161	244.9075	0.0498	0.0481
S2-F160	0.864	0.370	34.0	22.4	0.789	0.519	0.0200	0.0763	27.6683	275.3327	0.0444	0.0461
S2-F300	0.930	0.388	38.7	21.3	0.906	0.512	0.0117	0.0858	30.0014	247.1371	0.0370	0.0543
Outdoor Mean	0.919$\pm$0.030	0.324 $\pm$ 0.076	37.6$\pm$2.4	22.3 $\pm$ 3.0	0.902$\pm$0.049	0.506 $\pm$ 0.115	0.0137$\pm$0.0038	0.0808 $\pm$ 0.0229	28.4863$\pm$11.0698	268.3057 $\pm$ 20.3703	0.0399$\pm$0.0061	0.0467 $\pm$ 0.0037
<i>Indoor (6 scenes)</i>												
S3-F453	0.839	0.168	36.5	28.2	0.860	0.724	0.0150	0.0391	26.1754	304.6365	0.0418	0.0389
S4-F353	0.970	0.366	45.8	18.3	0.936	0.485	0.0051	0.1220	27.5573	262.6228	0.0489	0.0440
S5-F497	0.984	0.371	52.8	27.1	0.983	0.864	0.0023	0.0441	32.9267	272.8659	0.0459	0.0508
S5-F527	0.791	0.111	39.6	25.4	0.915	0.848	0.0105	0.0538	6.0017	293.5447	0.0431	0.0474
S6-F175	0.945	0.342	40.3	27.5	0.897	0.589	0.0096	0.0420	37.0109	265.2215	0.0378	0.0534
S6-F575	0.924	0.372	42.2	20.6	0.917	0.569	0.0077	0.0931	15.7710	269.3423	0.0368	0.0391
Indoor Mean	0.909$\pm$0.070	0.288 $\pm$ 0.107	42.9$\pm$5.2	24.5 $\pm$ 3.7	0.918$\pm$0.037	0.680 $\pm$ 0.143	0.0084$\pm$0.0040	0.0657 $\pm$ 0.0311	24.2405$\pm$10.4695	278.0390 $\pm$ 15.5574	0.0424$\pm$0.0042	0.0456 $\pm$ 0.0055
Overall Mean	0.914$\pm$0.053	0.307 $\pm$ 0.093	40.0$\pm$4.8	23.3 $\pm$ 3.5	0.910$\pm$0.044	0.586 $\pm$ 0.155	0.0112$\pm$0.0047	0.0738 $\pm$ 0.0280	26.5267$\pm$11.4986	272.7980 $\pm$ 18.9390	0.0410$\pm$0.0055	0.0462 $\pm$ 0.0046

4.1 Training Evaluation: ADC Rendering and Range–Azimuth

We compare rendered and measured range–azimuth (RA) maps on the cascaded radar used for optimization. Table 2 reports Pearson correlation, PSNR, SSIM, RMSE, and ADC-level magnitude metrics across all seven scenes, with Sionna-RT [18] as baseline. Just as the loss is defined on RA magnitude (Eq. 8), we evaluate in the RA domain because it implicitly validates phase accuracy. The azimuth FFT converts relative phase differences across virtual elements into angular information, so any systematic phase error due to per-element bias, incorrect path-length scaling, or incoherent phasor accumulation would smear or shift targets in the angular domain, degrading all RA metrics. Absolute phase is not meaningful in practice, as it depends on user-configurable initial phase, oscillator drift, and phase noise. These common-mode offsets appear in the DC term of the spatial FFT and are dropped during processing (Sec. B.3).

mmIR achieves a mean Pearson correlation of **0.919** on the seven outdoor scenes and **0.909** on six additional ColoRadar indoor scenes (lab, hallway, classroom), for an overall mean of **0.914**, substantially outperforming Sionna-RT (0.324/0.288/0.307 outdoor/indoor/overall). Table 2 quantifies this gap across all metrics, and Figs. 4 and 5 provide the qualitative comparisons, sorted by correlation. mmIR reproduces both the dominant scatterer locations and the diffuse multipath structure present in the ground truth, whereas Sionna-RT produces attenuated responses with spatially misaligned features.

Three architectural differences drive this gap. First, mmIR’s physically-based BSDF with six per-vertex learnable parameters captures spatially varying reflectance that Sionna-RT’s per-object scattering model cannot represent. This is most evident in geometrically complex scenes such as S1-F438 and S0-F135, where mixed-material surfaces dominate the radar response. Second, mmIR jointly optimizes surface normals and antenna beam patterns, parameters entirely outside Sionna-RT’s optimization scope, correcting both surface orientation errors inherited from noisy LiDAR tessellation and systematic antenna gain biases. Third, Sionna-RT’s path solver executes outside the AD graph. Ray tracing runs under

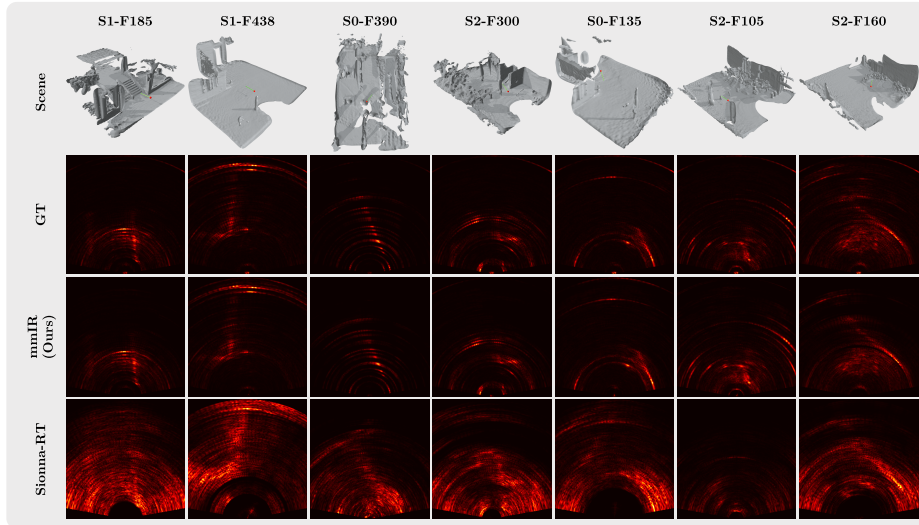


Fig. 4: Qualitative RA on the seven outdoor scenes, sorted by mmIR correlation. Rows: scene mesh (red sphere = radar, green arrow = boresight), GT, mmIR, Sionna-RT. mmIR captures dominant scatterers and multipath; Sionna-RT is weaker and misaligned.

gradient suspension and all path geometry is subsequently detached, so while material gradients propagate through field coefficients at each bounce, the path topology itself is frozen and no gradients reach vertex positions or normals, and a material change at one surface cannot redirect subsequent bounces. mmIR places the entire forward model inside a single computation graph, enabling end-to-end gradient flow from the loss through ADC synthesis, phasor accumulation, BSDF evaluation, and ray-triangle intersections. The weakest outdoor scene, S2-F160 (0.864), shows that even challenging geometries with complex multipath achieve high correlation under the min-max-normalized loss; residual gaps are dominated by sub-wavelength LiDAR-to-radar misalignment (Sec. B.2), to which coherent radar rendering is sensitive (Sec. D). mmIR substantially outperforms the baseline on every scene, indicating that the richer parameter space and deeper gradient flow consistently improve fit quality.

4.2 Radar Transfer Evaluation

A central claim of physics-based inverse rendering is that the optimized scene captures true physical properties, not sensor-specific patterns. We validate this by *cross-sensor transfer*: the scene parameters trained on the cascaded radar (12Tx×16Rx spanning 86 uniformly sampled virtual elements along azimuth) are frozen and used to render ADC for the single-chip radar (3Tx×4Rx, 12 virtual elements total spanning 8 uniformly sampled virtual elements along azimuth) that shares the same physical mounting. We apply the same rigid-body alignment transformation learned during cascaded training to the single-chip configuration

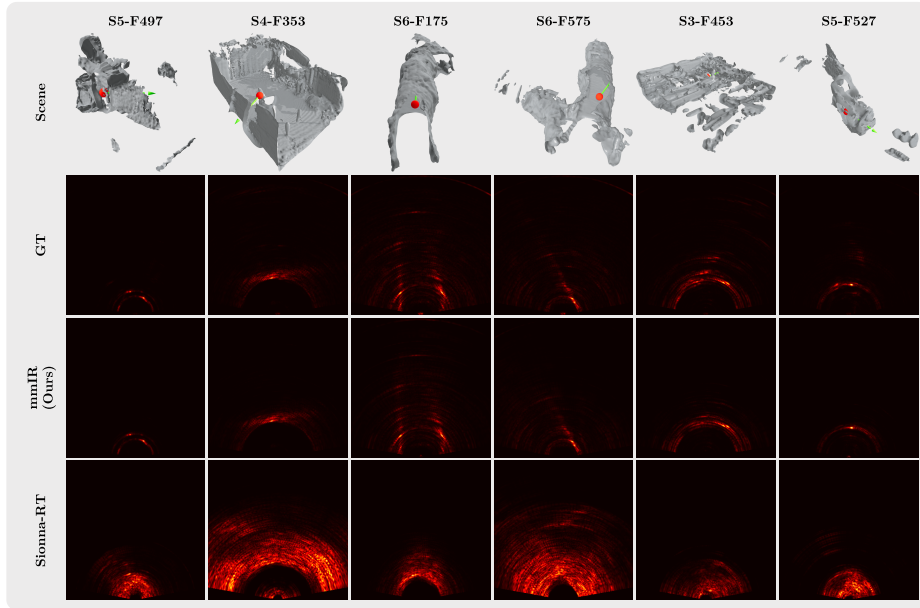


Fig. 5: As Fig. 4, for six indoor scenes (lab, hallway, classroom). mmIR tracks the GT returns; Sionna-RT over-responds in dense indoor multipath.

and render forward-only. In Table 3, mmIR achieves a mean transfer correlation of **0.554** on outdoor scenes, outperforming Sionna-RT (0.114). This indicates that mmIR has learned transferable scene physics rather than memorizing the training sensor’s idiosyncrasies, via optimized materials and normals that generalize across sensor configurations.

4.3 Inference: 3D Occupancy via Dense Virtual Array

After training, we freeze the optimized scene parameters and re-render from a dense virtual array (100Tx×100Rx, 10,000 virtual elements with $\lambda/2$ spacing) inspired by the Rohde & Schwarz QAR50 [52]. The cascade’s near-1D virtual aperture yields $256 \times 127 \times 1$ RAE resolution (Sec. A.1, A.3); only the dense array resolves true elevation at $256 \times 127 \times 127$ RAE resolution. We convert synthesized ADC into a 3D tensor via Range–Azimuth–Elevation (RAE) FFT and extract a sparse radar point cloud by thresholding at the 99.9th intensity percentile (Fig. 6). Following [27], we use single-bounce propagation at test time for the dense array.

We evaluate against LiDAR after radar-guided filtering: (1) filtering by the dense array’s 0 dB beam limits, (2) projecting LiDAR points onto the radar polar voxel grid, and (3) retaining only points where radar normalized intensity ≥ 0.1 . Occupancy metrics use KD-tree nearest neighbors with $\tau = 0.5$ m threshold. Geometric fidelity is measured via RMSE and Relative Chamfer

Table 3: Cross-sensor transfer evaluation. Parameters trained on the cascade (12Tx×16Rx) are frozen and used to render the single-chip AWR1843 (3Tx×4Rx) without re-training. Best in **bold**.

Scene	RA Correlation $\uparrow$		RA PSNR $\uparrow$		RA SSIM $\uparrow$		RA RMSE $\downarrow$		ADC Log-Mag MSE $\downarrow$		ADC MM-Mag MSE $\downarrow$	
	Ours	Bench.	Ours	Bench.	Ours	Bench.	Ours	Bench.	Ours	Bench.	Ours	Bench.
S0-F135	0.533	0.002	25.6	17.8	0.505	0.464	0.0524	0.1283	10.9639	301.4341	0.2194	0.0668
S0-F390	0.464	0.328	21.2	24.2	0.529	0.566	0.0873	0.0619	12.9644	331.2549	0.0630	0.0840
S1-F185	0.567	0.136	22.4	20.1	0.422	0.480	0.0759	0.0990	0.5754	324.8441	0.1993	0.0771
S1-F438	0.684	0.079	23.2	13.5	0.674	0.202	0.0694	0.2119	21.8672	348.0237	0.1507	0.0662
S2-F105	0.608	0.084	24.7	22.8	0.609	0.690	0.0583	0.0721	3.4458	289.7929	0.1959	0.1068
S2-F160	0.276	0.159	19.3	18.3	0.493	0.463	0.1085	0.1219	20.1530	338.1699	0.0893	0.0798
S2-F300	0.745	0.008	27.2	23.6	0.468	0.412	0.0434	0.0661	7.7226	287.9379	0.1074	0.0948
Mean	0.554±0.143	0.114±0.103	23.4±2.5	20.0±3.6	0.529±0.080	0.468±0.138	0.0707±0.0206	0.1087±0.0488	11.0989±7.3886	317.3511±22.3659	0.1464±0.0565	0.0822±0.0136

Table 4: 3D occupancy vs. GT LiDAR. Baseline: ColoRadar’s cascaded-radar point cloud with elevation extrusion added in post-processing. Ours: mmIR’s RAE rendered via the dense virtual aperture (100×100 elements), resolving true elevation. KD-tree match at $\tau=0.5$ m. Best in **bold**.

Scene	Precision $\uparrow$		Recall $\uparrow$		F1 $\uparrow$		Accuracy $\uparrow$		RMSE $\downarrow$		R-CD $\downarrow$	
	Ours	ColoR	Ours	ColoR	Ours	ColoR	Ours	ColoR	Ours	ColoR	Ours	ColoR
S0-F135	0.077	0.223	0.018	0.054	0.029	0.087	0.044	0.057	2.324	3.302	0.248	0.228
S0-F390	0.481	0.045	0.110	0.014	0.179	0.021	0.212	0.017	2.692	2.745	0.216	0.207
S1-F185	0.997	0.483	0.566	0.082	0.722	0.141	0.703	0.103	0.909	2.380	0.067	0.236
S1-F438	0.633	0.433	0.265	0.117	0.373	0.185	0.362	0.131	1.511	3.339	0.116	0.268
S2-F105	0.771	0.154	0.162	0.025	0.268	0.044	0.340	0.029	1.099	3.402	0.113	0.282
S2-F160	0.516	0.383	0.177	0.165	0.264	0.231	0.257	0.183	1.917	1.611	0.146	0.118
S2-F300	0.850	0.508	0.501	0.122	0.630	0.197	0.601	0.130	1.040	1.643	0.083	0.118
Mean	0.618±0.278	0.319±0.165	0.257±0.189	0.083±0.051	0.352±0.228	0.129±0.075	0.360±0.210	0.093±0.056	1.642±0.638	2.632±0.721	0.141±0.062	0.208±0.061

Distance (R-CD). The geometric input is a smoothed Poisson mesh aggregated from ~ 50 LiDAR sweeps, the evaluation ground truth is a single raw LiDAR frame at a different time instant, and mmIR’s output is the radar-physics-gated 99.9th-percentile RAE peak set, not the input mesh (Sec. B.6). Table 4 shows substantial variation across outdoor scenes driven by radar visibility. Scenes with multiple surfaces within the field-of-view achieve strong reconstruction (S1-F185: Precision=0.997, R-CD=0.067), while open scenes with ground planes nearly parallel to boresight show degraded recall as specular energy is directed away from receivers. Component ablations (Sec. C.4) confirm each module contributes (e.g., -3.7% correlation without antenna optimization, -1.0% without diffraction). The supplement details training and implementation hyperparameters (Sec. B.4), runtime (Sec. C.5), dense-array inference and 3D-occupancy post-processing (Sec. B.5, B.6), and additional qualitative results and ablations (Sec. C).

5 Conclusion

We introduced mmIR, an end-to-end differentiable inverse renderer that fits per-vertex ITU materials, normals, and antenna patterns from raw FMCW ADC via automatic differentiation through a phase-coherent MIMO forward model. On seven outdoor and six indoor scenes, mmIR achieves 0.914 mean Pearson correlation versus 0.307 for Sionna-RT on rendered RA images. We further validate cross-sensor generalization by transferring trained scenes to a different radar configuration. We then show that mmIR’s fitted scenes re-render dense virtual-aperture ADC for 3D occupancy detection and scene reconstruction,

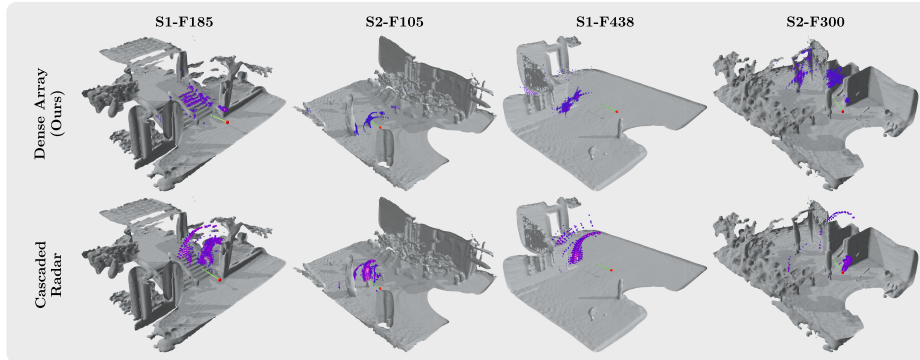


Fig. 6: Single-frame 3D occupancy on the LiDAR mesh. Top: mmIR’s rendered RAE via the dense virtual aperture (100×100 elements), thresholded at the 99.9th intensity percentile. Bottom: ColoRadar-provided RAE, obtained from the near-1D cascaded radar with elevation extrusion added in post-processing. Red sphere = radar, green arrow = boresight.

validated against LiDAR. This opens a path toward physics-grounded radar sensing pipelines, including calibration, radar-aware 3D reconstruction, and joint multi-modal optimization [48].

Limitations and future work. mmIR’s rendered ADC is bottlenecked by mesh quality, and the inverse renderer is sensitive to LiDAR-radar alignment for both cascade and single-chip configurations. Our method supports surface-only interactions and single-bounce diffraction, and does not model transmission, refraction, or volumetric scattering. We do not account for Doppler phase from ego-motion, which would require a TDM-compatible $4\pi \mathbf{v}_{\text{ego}} \cdot \mathbf{t} / \lambda$ term, and assume static scenes, leaving dynamic-scene modeling for future work. Vertex-position and pose gradients are supported, but we found that at 77 GHz ($\lambda \approx 3.9$ mm) sub-millimeter shifts induce large phase changes that destabilize these gradients under current Monte Carlo budgets and are therefore disabled. True shape-from-radar remains compelling but will likely require dense multi-view supervision, regularization strategies and priors beyond our single-frame RA fitting setting. Evaluation is also confined to ColoRadar’s clear-weather captures, with adverse weather evaluation left to future work. Validating the inferred materials against ground-truth dielectric properties, for instance through controlled anechoic-chamber captures of known materials, is another promising direction, since ColoRadar provides no co-located RGB or ground-truth materials for this comparison. See the supplement (Sec. D) for details.

References

1. Bangaru, S.P., Gharbi, M., Luan, F., Li, T.M., Sunkavalli, K., Hasan, M., Bi, S., Xu, Z., Bernstein, G., Durand, F.: Differentiable rendering of neural sdfs through reparameterization. In: SIGGRAPH Asia 2022 Conference Papers. SA ’22,

- Association for Computing Machinery, New York, NY, USA (2022). <https://doi.org/10.1145/3550469.3555397>, <https://doi.org/10.1145/3550469.3555397>
2. Bangaru, S.P., Li, T.M., Durand, F.: Unbiased warped-area sampling for differentiable rendering. *ACM Trans. Graph.* **39**(6) (Nov 2020). <https://doi.org/10.1145/3414685.3417833>, <https://doi.org/10.1145/3414685.3417833>
 3. Bangaru, S.P., Wu, L., Li, T.M., Munkberg, J., Bernstein, G., Ragan-Kelley, J., Durand, F., Lefohn, A., He, Y.: Slang.d: Fast, modular and differentiable shader programming. *ACM Trans. Graph.* **42**(6) (Dec 2023). <https://doi.org/10.1145/3618353>, <https://doi.org/10.1145/3618353>
 4. Barnes, D., Gadd, M., Murcutt, P., Newman, P., Posner, I.: The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). pp. 6433–6438 (2020). <https://doi.org/10.1109/ICRA40945.2020.9196884>
 5. Beckmann, P., Spizzichino, A.: The scattering of electromagnetic waves from rough surfaces. Norwood (1987)
 6. Bitterli, B., Wyman, C., Pharr, M., Shirley, P., Lefohn, A., Jarosz, W.: Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting. *ACM Trans. Graph.* **39**(4) (Aug 2020). <https://doi.org/10.1145/3386569.3392481>, <https://doi.org/10.1145/3386569.3392481>
 7. Born, M., Wolf, E.: Principles of optics: electromagnetic theory of propagation, interference and diffraction of light. Elsevier (2013)
 8. Borts, D., Liang, E., Broedermann, T., Ramazzina, A., Walz, S., Palladin, E., Sun, J., Brueggemann, D., Sakaridis, C., Van Gool, L., Bijelic, M., Heide, F.: Radar fields: Frequency-space neural scene representations for fmcw radar. In: ACM SIGGRAPH 2024 Conference Papers. SIGGRAPH '24, Association for Computing Machinery, New York, NY, USA (2024). <https://doi.org/10.1145/3641519.3657510>, <https://doi.org/10.1145/3641519.3657510>
 9. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuscenes: A multimodal dataset for autonomous driving. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (June 2020)
 10. Chen, X., Ding, J., Zheng, K., Fang, X., Zhang, X., Lu, C.X., Li, Z.: Invertwin: Solving inverse problems via differentiable radio frequency digital twin (2025), <https://arxiv.org/abs/2508.14204>
 11. Deng, J., Xie, P., Zhang, L., Cong, Y.: Isar-nerf: Neural radiance fields for 3-d imaging of space target from multiview isar images. *IEEE Sensors Journal* **24**(7), 11705–11722 (2024). <https://doi.org/10.1109/JSEN.2024.3360981>
 12. Dodds, L., Boroushaki, T., Ham, C., Adib, F.: Mito: A millimeter-wave dataset and simulator for non-line-of-sight perception (2025), <https://arxiv.org/abs/2502.10259>
 13. Dudek, M., Wahl, R., Kissinger, D., Weigel, R., Fischer, G.: Millimeter wave fmcw radar system simulations including a 3d ray tracing channel simulator. In: 2010 Asia-Pacific Microwave Conference. pp. 1665–1668 (2010)
 14. Gao, X., Xing, G., Roy, S., Liu, H.: Experiments with mmwave automotive radar test-bed. In: 2019 53rd Asilomar Conference on Signals, Systems, and Computers. pp. 1–6 (2019). <https://doi.org/10.1109/IEEECONF44664.2019.9048939>
 15. Gargallo, P., Prados, E., Sturm, P.: Minimizing the reprojection error in surface reconstruction from images. In: 2007 IEEE 11th International Conference on Computer Vision. pp. 1–8 (2007). <https://doi.org/10.1109/ICCV.2007.4409003>

16. Guan, J., Madani, S., Jog, S., Gupta, S., Hassanieh, H.: Through fog high-resolution imaging using millimeter wave radar. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 11461–11470 (2020). <https://doi.org/10.1109/CVPR42600.2020.01148>
17. Hofmann, N., Wirth, V., Bräunig, J., Ullmann, I., Vossiek, M., Weyrich, T., Stamminger, M.: Inverse rendering of near-field mmwave mimo radar for material reconstruction. *IEEE Journal of Microwaves* **5**(2), 356–372 (2025). <https://doi.org/10.1109/JMW.2025.3535077>
18. Hoydis, J., Aoudia, F.A., Cammerer, S., Nimier-David, M., Binder, N., Marcus, G., Keller, A.: Sionna rt: Differentiable ray tracing for radio propagation modeling. In: 2023 IEEE Globecom Workshops (GC Wkshps). pp. 317–321 (2023). <https://doi.org/10.1109/GCWkshps58843.2023.10465179>
19. Huang, T., Miller, J., Prabhakara, A., Jin, T., Laroia, T., Kolter, Z., Rowe, A.: Dart: Implicit doppler tomography for radar novel view synthesis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 24118–24129 (June 2024)
20. ITU-R: Effects of building materials and structures on radiowave propagation above about 100 MHz. Recommendation P.2040-3, International Telecommunication Union (2023)
21. Jakob, W., Speierer, S., Roussel, N., Vicini, D.: Dr.jit: a just-in-time compiler for differentiable rendering. *ACM Trans. Graph.* **41**(4) (Jul 2022). <https://doi.org/10.1145/3528223.3530099>, <https://doi.org/10.1145/3528223.3530099>
22. Jin, X., Xu, C., Gao, R., Wu, J., Wang, G., Li, S.: Diffsound: Differentiable modal sound rendering and inverse rendering for diverse inference tasks. In: ACM SIGGRAPH 2024 Conference Papers. SIGGRAPH '24, Association for Computing Machinery, New York, NY, USA (2024). <https://doi.org/10.1145/3641519.3657493>, <https://doi.org/10.1145/3641519.3657493>
23. Kato, H., Beker, D., Morariu, M., Ando, T., Matsuoka, T., Kehl, W., Gaidon, A.: Differentiable rendering: A survey (2020), <https://arxiv.org/abs/2006.12057>
24. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: Proceedings of the Fourth Eurographics Symposium on Geometry Processing. p. 61–70. SGP '06, Eurographics Association, Goslar, DEU (2006)
25. Kerbl, B., Kopanas, G., Leimkuehler, T., Drettakis, G.: 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* **42**(4) (Jul 2023). <https://doi.org/10.1145/3592433>, <https://doi.org/10.1145/3592433>
26. Kramer, A., Harlow, K., Williams, C., Heckman, C.: Coloradar: The direct 3d millimeter wave radar dataset. *The International Journal of Robotics Research* **41**(4), 351–360 (2022). <https://doi.org/10.1177/02783649211068535>, <https://doi.org/10.1177/02783649211068535>
27. Kung, P.C., Harisha, S., Vasudevan, R., Eid, A., Skinner, K.A.: Radarsplat: Radar gaussian splatting for high-fidelity data synthesis and 3d reconstruction of autonomous driving scenes. In: 2025 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 27596–27606 (2025). <https://doi.org/10.1109/ICCV51701.2025.02562>
28. Lei, Z., Xu, F., Wei, J., Cai, F., Wang, F., Jin, Y.Q.: Sar-nerf: Neural radiance fields for synthetic aperture radar multi-view representation (2023), <https://arxiv.org/abs/2307.05087>
29. Li, T.M., Aittala, M., Durand, F., Lehtinen, J.: Differentiable monte carlo ray tracing through edge sampling. *ACM Trans. Graph.* **37**(6) (Dec 2018). <https://doi.org/10.1145/3272127.3275109>, <https://doi.org/10.1145/3272127.3275109>

30. Li, X., Wang, X., Yang, Q., Fu, S.: Signal processing for tdm mimo fmcw millimeter-wave radar sensors. *IEEE Access* **9**, 167959–167971 (2021). <https://doi.org/10.1109/ACCESS.2021.3137387>
31. Lin, B., Li, C., Ji, Y., Liu, X., Fang, G.: A millimeter-wave 3d imaging algorithm for mimo synthetic aperture radar. *Sensors* **23**(13) (2023). <https://doi.org/10.3390/s23135979>, <https://www.mdpi.com/1424-8220/23/13/5979>
32. Lin, D., Kettunen, M., Bitterli, B., Pantaleoni, J., Yuksel, C., Wyman, C.: Generalized resampled importance sampling: foundations of restrir. *ACM Trans. Graph.* **41**(4) (Jul 2022). <https://doi.org/10.1145/3528223.3530158>, <https://doi.org/10.1145/3528223.3530158>
33. Loubet, G., Holzschuch, N., Jakob, W.: Reparameterizing discontinuous integrands for differentiable rendering. *ACM Trans. Graph.* **38**(6) (Nov 2019). <https://doi.org/10.1145/3355089.3356510>, <https://doi.org/10.1145/3355089.3356510>
34. Lu, H., Vattheuer, C., Mirzsoleiman, B., Abari, O.: Newrf: A deep learning framework for wireless radiation field reconstruction and channel prediction (2024), <https://arxiv.org/abs/2403.03241>
35. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: representing scenes as neural radiance fields for view synthesis. *Commun. ACM* **65**(1), 99–106 (Dec 2021). <https://doi.org/10.1145/3503250>, <https://doi.org/10.1145/3503250>
36. Mostajabi, M., Wang, C.M., Ranjan, D., Hsyu, G.: High-resolution radar dataset for semi-supervised learning of dynamic objects. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (June 2020)
37. Nimier-David, M., Vicini, D., Zeltner, T., Jakob, W.: Mitsuba 2: a retargetable forward and inverse renderer. *ACM Trans. Graph.* **38**(6) (Nov 2019). <https://doi.org/10.1145/3355089.3356498>, <https://doi.org/10.1145/3355089.3356498>
38. Rafidashti, M., Lan, J., Fatemi, M., Fu, J., Hammarstrand, L., Svensson, L.: Neuradar: Neural radiance fields for automotive radar point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. pp. 2513–2523 (June 2025)
39. Ramamoorthi, R., Mahajan, D., Belhumeur, P.: A first-order analysis of lighting, shading, and shadows. *ACM Trans. Graph.* **26**(1), 2–es (Jan 2007). <https://doi.org/10.1145/1189762.1189764>, <https://doi.org/10.1145/1189762.1189764>
40. Schüßler, C., Hoffmann, M., Bräunig, J., Ullmann, I., Ebelt, R., Vossiek, M.: A realistic radar ray tracing simulator for large mimo-arrays in automotive environments. *IEEE Journal of Microwaves* **1**(4), 962–974 (2021). <https://doi.org/10.1109/JMW.2021.3104722>
41. Steinberg, S., Ramamoorthi, R., Bitterli, B., Mollazainali, A., D’Eon, E., Pharr, M.: A free-space diffraction bsdf. *ACM Trans. Graph.* **43**(4) (Jul 2024). <https://doi.org/10.1145/3658166>, <https://doi.org/10.1145/3658166>
42. Sun, S., Petropulu, A.P., Poor, H.V.: Mimo radar for advanced driver-assistance systems and autonomous driving: Advantages and challenges. *IEEE Signal Processing Magazine* **37**(4), 98–117 (2020). <https://doi.org/10.1109/MSP.2020.2978507>
43. Takawale, H., Roy, N.: Spinr: Neural volumetric reconstruction for fmcw radars (2025), <https://arxiv.org/abs/2503.23313>
44. Talbot, J., Cline, D., Egbert, P.: Importance Resampling for Global Illumination. In: Bala, K., Dutre, P. (eds.) *Eurographics Symposium on Rendering* (2005). The Eurographics Association (2005). <https://doi.org/10.2312/EGWR/EGSR05/139-146>

45. Ulaby, F.T., Ravaioli, U.: Fundamentals of applied electromagnetics, vol. 7. Pearson Upper Saddle River, NJ (2015)
46. Veach, E.: Robust Monte Carlo methods for light transport simulation. Stanford University (1998)
47. Veach, E., Guibas, L.J.: Optimally combining sampling techniques for monte carlo rendering. In: Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques. p. 419–428. SIGGRAPH '95, Association for Computing Machinery, New York, NY, USA (1995). <https://doi.org/10.1145/218380.218498>, <https://doi.org/10.1145/218380.218498>
48. Vilesov, A., Chari, P., Armouti, A., Harish, A.B., Kulkarni, K., Deoghare, A., Jalilian, L., Kadambi, A.: Blending camera and 77 ghz radar sensing for equitable, robust plethysmography. *ACM Trans. Graph.* **41**(4) (Jul 2022). <https://doi.org/10.1145/3528223.3530161>, <https://doi.org/10.1145/3528223.3530161>
49. Walter, B., Marschner, S.R., Li, H., Torrance, K.E.: Microfacet Models for Refraction through Rough Surfaces. In: Kautz, J., Pattanaik, S. (eds.) *Rendering Techniques*. The Eurographics Association (2007). <https://doi.org/10.2312/EGWR/EGSR07/195-206>
50. Watts, C.M., Lancaster, P., Pedross-Engel, A., Smith, J.R., Reynolds, M.S.: 2d and 3d millimeter-wave synthetic aperture radar imaging on a pr2 platform. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 4304–4310 (2016). <https://doi.org/10.1109/IROS.2016.7759633>
51. Wei, J., Luomei, Y., Zhang, X., Xu, F.: Learning surface scattering parameters from sar images using differentiable ray tracing. *IEEE Transactions on Geoscience and Remote Sensing* **62**, 1–15 (2024). <https://doi.org/10.1109/TGRS.2024.3459620>
52. Wirth, V., Bräunig, J., Hofmann, N., Vossiek, M., Weyrich, T., Stamminger, M.: Maroon: A dataset for the joint characterization of near-field high-resolution radio-frequency and optical depth imaging techniques. *ACM Trans. Graph.* **45**(2) (Mar 2026). <https://doi.org/10.1145/3796224>, <https://doi.org/10.1145/3796224>
53. Zeltner, T., Georgiev, I., Jakob, W.: Specular manifold sampling for rendering high-frequency caustics and glints. *ACM Trans. Graph.* **39**(4) (Aug 2020). <https://doi.org/10.1145/3386569.3392408>, <https://doi.org/10.1145/3386569.3392408>
54. Zhang, J., Li, Z., Wang, C., Wen, B.: Rf4d:neural radar fields for novel view synthesis in outdoor dynamic scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 15387–15397 (June 2026)
55. Zhang, K., Luan, F., Li, Z., Snavely, N.: Iron: Inverse rendering by optimizing neural sdfs and materials from photometric images. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 5555–5564 (2022). <https://doi.org/10.1109/CVPR52688.2022.00548>
56. Zhang, R., Cao, S.: 3d imaging millimeter wave circular synthetic aperture radar. *Sensors* **17**(6) (2017). <https://doi.org/10.3390/s17061419>, <https://www.mdpi.com/1424-8220/17/6/1419>
57. Zhang, Z., Roussel, N., Jakob, W.: Projective sampling for differentiable rendering of geometry. *ACM Trans. Graph.* **42**(6) (Dec 2023). <https://doi.org/10.1145/3618385>, <https://doi.org/10.1145/3618385>